

















pS1 (SEQ ID NO:1)

- 1 GATACGTAAA CAACGTGTAT CCAGTAAGTA TCAAGCCTAA TCTCGAAGTG
- 51 GTAGAGACGA GATGACCGAC TTTGACACTC ACTATGGGTC AATAATAATA
- 101 ACTGAAATAA AACTAAGATA TTTAAACCAA CATGATTTAC AGAATTTACA
 151 ATAATTTATT TAATCAGCAG AAATAATCAA ATTTCTTCAA ATGTAACAAT
- 201 TCTCAATATA TTAATTAAAT TCCTTCAATT CAAATAATTT CTAATTTATC
- 201 ICTCAATATA TAATTAAAA TICCTICAATT CAAATAATTI CTAATTATC
- 251 AATTAAACCT CATTTACAGG AGTAACAATT AATTCCTTAA CAAGCAAGAA
 301 TAATAATTCA TTAAATTCCA AGGATTTTTC AATTTATTAA TTAGCTTCAC
- 351 AACCTGAAAT AAATTATTAA AGTATCGTGT AATTATTATT ATTAAGCACG
- 351 AACCIGAAAT AAATTATTAA AGTATCGIGI AATTATTATT ATTAAGCAC
- 401 ATTTCTGCCG AGGACATACG GCCCGATCCA GAGTATC

pS4 (SEQ ID NO:2)

- 1 GATACTAGAG TGGTGTTATC AATTCTTACT CGTATGAATT AATTAAATTT
- 51 GTCTCTTATT TCTGTCCTAA GTCATATACA AGAAATGCTA ACTCCATCCG
- 101 TTTCAATCCC TATGACATAG TTTGATTTGA TTGAATTTGA AAATTTAAGA
- 151 AACAAAAGAT AATTTTTGTG ACTCATAATT TAGACATGTG TTATAAGACT
- 201 TTTCTCATGA ATTTTTTAGA AACAAATGAT AATTTTT<u>GGA ACTCATAATT</u>
- 251 TAGACGTTTT ATAAAAAATA CTAACTGCAT CTGGTTCAAT ATTTATGTGT
- 301 TATTCCTATA AAACTTCTGG ACTTATATTT TTAAATATTT CATAATATTT
- 351 GGTATCGGTA TAATTTTTTT GTCACTTTTG GATGAAAGGG AAGTTTAAGT
- 401 AAATTTCTTT TTCCAAATTT AGAAAGTTAT AATATTCTTT TTAAAACGCC
- 451 CAAAAAGAAA AATAAGCTAT TGATTATTAT AAGCCTAAAC CAAAAGAATT
- 501 CTTTGACTAG TAGGAAGCCA TTTTTAAGTT AGGCGCCAAA ATTCAAAGCC
- 551 AACGTGGGCA TATCTCCAAA CTGGCGGCTA CAGTATC

pS8 (SEQ ID NO:3)

- 1 ACCGCTTTTA TTATTATTAT TTTTACCGAG AATTACAACA TCATGAAAAT
- 51 ACATCTCGAA CCACGTCACA TCAATGCACC CGCGGTTATT GACATATTTC
- 101 AACTCTGTTG AGATTTGGAT TTGGGTCACA TAAATGTGCA CCCGAGTTTA
- 151 AGAGGATAAC ATTATTAAAT ACGCGCCTAA AACGACTAGC GTATCATTAT
 201 TTTGGGTAGG GCCGTGAAAT TTTGCTAAAC TGCCCATCCA GAAATCTAAG
- 251 TAATTITACC AACACGTATA GAGGGCCCCA CAGCTTGTGT ATTITTGTTT
- 251 FAATTITACC AACACGTATA GAGGGCCCCA CAGCTTGTGT ATTITTGTTT
 301 GTCGAGGCTC GTCTCATTCA TTATTTTTAA AAGGAATTTG CAACGTCGTG
- 351 GAAATGCATC TCGAACCACG TCACAATCAA TGA

pS115 (SEQ ID NO:4)

1 GAATTCGATA GACTCACTTA AATATTAGAA GTGAATTACC TAGAGTTAGA 51 TCCAAAACAA TTATCTTGCA CCTATCCTAT CAACCCTTAT CTTTTCCCAT 101 TGATTACTAC CTTGCTTACC TTTGTTACGA TTTTCATTAG ACAATAACTT 151 TAGATTCTTA GTTAATTGCA GTTAGAAATT ATATTAAATT TCAATTGTTG 201 GATCATCTTG AATACCAATC AAGCTAGAAA ATACAAGAAT ACTGTTTAAA 251 TCAAATCCAT GTGGATACGA TATTATACTA TATTATATTT GACTTGTGAG 301 CATTATTTAT GTGTGTTTTG TGCTCGTCAA AGTTTGGCGT CGTTGCCGAG 351 GATTGGCAAT CAATAGTGTT TGAAATAGTT TTTGGTGCTA ATTTAGGAAT 401 TAGGTTTTAT TTATTTATTT TITCTTTTCT TTTCTTTTC CTTTTCTATT 451 TTATTTCCTT TATTAGTTAA CTTCTTTTCA AGATTTTTTT TGTAGTACCT 501 AACAAGTTAG AGAAGATACT GTAGATTTTG AACTCTAAAT GTTGTGAAGA 551 TGGAGTACAA CCAGCCTAAG AAAATATTTG AATAGTTAGC AGCTGAACAT 601 TATCGGCGGT CGGTTATGCG GTTTAAATGC GGTGGAAGCA TCTACCACCG 651 CAGCCTAAAG AAAATATTTT GAATAGTTAG CAGCTTGAAC ATTATCGGCG 701 GTCGGTTATG TGTTTTAAAT GCGGTGGAAA TCATCTACGG GCTAACTGTC 751 AAGCAGGTAT GTATTCTTCC TATGGTTCGT ATTTTGAGGA GTCTCACTCT 801 GTTTCTAGTT CGTACATGTA TGAGGATTCA TATGGGCACA ACTCTGACTC 851 TGGTTGGGAT GAATTC

pS116-1.1B (SEQ ID NO:5)

1 GAATTGTATT ATTGTTAGGT GGGAGAGATT TTTGACTATA TGGGTTAAAA 51 TCAGCGACAA AGGGCCAAAT ATACCTATTT ACTTTTAAAA ATAGTCTAAT 101 AATACCTCTC GTTATATTAT TAGGTTATCT ATACCTTTGC AGTCATATTT 151 TGGGTTCAAA TATACCCCTC ATTTAAACGG AGGGACACGT GTCATCGTCC 201 TGTTGGTCAA TTCTAAATAT CTCCTAATTA ATTAAAAAGA CTCATTACCC 251 ATATCCGAAA AATATTTTTT AAAGCAATAT TTTTTTATAA AAAATGGAAA 301 AACTGAAATT ATTTTTACTA AAAATTGAAA AAAACGAAAA TAGTTTTTTT 351 TCAGTTTTTA CAAAAAAACT ATTTTAGAAA AAATTGAAAA ATATTTTCTA 401 AAACAATGTT TTTGTAAAAA CTGAAAAAAA AGAAGCTGAA AATCAATTTT 451 CTAAAGCAAT TTTATTTGTA AAATCTGGAA AAAACTACTA AAAACTGAAA 501 AAATGAAAAT ATTTTTTTT CTAATTTTTA CAAAAAAAAC TGCTTTAAAA 551 AAAGCTGAAA ATATTTTCTA AAACAATATT TITGTAAAAA CTAAAAAAAA 601 AATATTTTCT TCTTTTTTC AGTTTTTAGT TAAAAATATT TAAGTTTTTT 651 CCAGTITTTA ATTACTITAG AAAATTACTT TTCTGCTTTT TTTTCAGTTT 701 TTACAAAAAT ATTATTTTAG AAAATATTTT TCAGTTCTTT AAAGCAGTTT 751 TTTTTTGTAA AAACTTGAAA AACAATATTT TCGTTTTTTT CAGTTTTTAG 801 TAAAATTTGT TTTTAGTTTT TTTCAGTTTT TACCAAAAAT AAAATTGCTT 851 TAGAAAATTA TTTTTCGGGT ATGGGTAATG GGTCTTTTTA ATTAATTAGG 901 AGATATTTTG AATTGATCAA TAGGACGATG ACACATGTCC CTCCGTTTAA 951 ATGAGGTGTA TATTTGAATC CAAAGTAAGA CTGCAGCCCG GGGGATCC

pS202-1 (SEQ ID NO:6)

1 GAATTICGATA TGGCTTGTTG GACAAGAATT AATGAATCAA TIGTGAAAAA
51 GTTGATGGAC ATATTGAAGG TAAAATCATA TACTATTTTT CTAAAATCTC
101 TITTAAATGT TCCCCAATTA TCTGATTTCT ATATTGCTCT TAAATGTCAC
151 TCAACCTTAG ATCAACAAAC ATATAACTTA CCCAGTACAT AAGAGATTGC
201 GGCATTATGG CTTGAAGAAA ATCCTAGAGA CACATCTGCA CCACATATTT
251 GAATTIATAC CCACAGTAAT AGAGCTCGGT TAGTACATTA TTATTATGGA
301 TGTTACGATC CGTTGCAGTA TCCATTATTA TITTCCTTCG GTGAAATGG
301 ATGACATTGT GGAATTAAAA AAATTATTCA GACAAAAAAT TCGACGAAAC
401 GTAGAGCTTA CTGCGAACAT GAACAATGC CCAGTATATC AAATACGTGT
451 TCAGTTGATG GATTCCTTGA TATGGAAGAT GAATCACTAC AAACAGGAAA
501 ACGAAAAAAGA GATACAGTGT CTTGTCGAGA GTATTATTGT TACAAATTC
601 CAACAATTTA TAGTAGATAT ATATATAA AGCTT

pS202-2 (SEQ ID NO:7)

1 AAGCTTGCAC GCCTACATCG TGGGATAATT TAGAAAAAGG AAAGGGTATA 51 TTGGATCCCC CTATCATTTG TGAAACAGGT AACCATACGA GAACCCCTTT 101 CGCTTCCTGA AAAATGTTAT ATATTGTTGT ACTCATATTT ATACACTATT 151 TATTATTAAT ATAACGATGC TTATTTTGCT TGGAGATTGG AGATTATCAC 201 AGCTTATTTA TCTTATATTG TATCTTATTA AACTTAAAAA CATAAATACT 251 ACGTGCTCTT TTAATTTGGG ATCTATTAAG GGTTCGTTGC ACGCTTTTAA 301 ACATCTTGGC TATTCTGTTT ACCAGCTGCT ACCTTAGCCT GTATGCTTAC 351 ATCATCTCCT AATTTAGACA AAGGAAAGGG TATATTGGAC CCCCCCTATC 401 ATTCGTGAAA CAGGTAAACA TACATTCAGA TTATACTCTT TTCAGAATGA 451 CATATTGTTT ATACATTACT GTAAATTGTG ACTATTTGTA TATTAGGGTC 501 CACATCGGGT ACATCTAACC TGCGTCATGT TATCTTGAAC ACTGTTCCAA 551 TCAAAGGTTT GCACAAACTT AATGTTACAA TCATGTCCAC CATACGTATG 601 CCTTGGTGCT CTTTTTTTC CTAATGATAC TTCTTATATA TTCAGCTCAT 651 AGGCGGGCCA GAAAGGTGTG CCTGGTCACT AAAGAGCAAC GAAGTGAGTA 701 TGTTGCTCTA AAAAGGGTCC CACACTGTCA ATTCTGTCAT CCAAAGAAGT 751 TTGAATATGA ACCTCCAGGA TTTTGCTGTA ACAGTGGTTC AATAAGGTTG 801 ACATCTCATA AAATGCCAAC TGAATTATCG GAGTTATACT TTGGAAATAC 851 TGAAGAATCT GAAAATTTTC GAACTTATAT TAGAACATAC AATAACATGT 901 TTGCATTTAC TTCACTTGGT GTCAAGTATG ATAAAGAGCT AGCGAGAAGA 951 AATTGTGGTA TCTACACATT TAGAGTCCAG GGACAGATGT ATCATTITAT 1001 AGATGATTTA GTTCCTTCCA ATGAAAAACC TAGGAATTTA TAGCTGTACT 1051 TCTACGATAA TGATAATGAA CTAGCCAAAT CAAGCTT

pS205-2 (SEO ID NO:8)

1 GAATTCTTCA GCCATTGTAC ATATAGTTGT GTATTAATGT TATTAATAAT 51 GGATAATTAA ATATACCT GGAATAAATA TACGATATTA TAATAGTGTG 101 TAATTATATA TAAAAATTAT ACATAATATA ATGATGGTAT TTAATATAGC 151 ATAAATTTGA ACGATCTGGA TTGATTTCTT GAATCAAAAT AGAGTTGTGT 201 GAAAAGAAA GAATGAGATG AAAAGCAAAG TATGAAGAGA TGAATTTGTG 251 TTTTTTTAT GGAGGAGGAA GGTTCTCAGT GATGGAATCA TCCCTGGTTT 351 AAAAAAAGG GAGAGAGAGT AGAATGGAAC GGCTAGGTGA AAGTATAGGA 401 GTAGAAATTA GGTTCAGGGA GAGAAAAGGG GGGAAATTAA TTCCTAAATT 451 AATGGGATTC TAATTTITAA ACTGTTTTGA AATATTITAA AAGTAGTGTT 501 ATTTATATTA TTAACTTTTA AAAAAAGTCA AACGAGGTAA AAATTCCATG 551 GGGGAAAATT TAAATGGTTA GTCTTCTATA ATATTTTCAA CTCTGCTTAG 601 CACTAAAAAT TAGTCTAAAA ATAACCCTAA ATTAGTGTAT CTAAATTAAT 651 TAGTTCATCG AACAGGAGCA TTGGATTATC CCTCCAGAGT TACACAGGAA 701 GCTT

pS206-1 (SEQ ID NO:9)

1 GGATCCAGCT ATTATTATAG CATGTGAGTT GTCCGTGAAC AGCTAATTTT 51 TTACCACACC CAAATTCAAT ACTATTTTAG TGTAAATATA TCTTTTAGGT 101 CTAGTCTTAA TATTTAACTT TITGTCTTAC TTTTAATAGA TITTATTTGA 151 GAAAAATTAA TAATTACAAA AAATAAAAAG TATATATTCA CATACTTATA 201 GTACAAACTT TGTTTCTATT TATAAAGAGA AAAAGAAATT TTACAAAAAA 251 CAAATATATT TGCTTTCTTT TAATTAGTAG TTTTATTAAG CAAGCTATAG 301 AAGCTC

pS211-1 (SEO ID NO:10)

1 GAATTCCGTG GTTTTAGCAC GGTCGCTCAA TTGTCATATT TGGCTCATTT 51 ATCTGATTTT TAAACAATTA AGAACTTATA TGCAAATTTA ACTTTTAAAA 101 CCGCTTTTAT CATTATTTAT TTTATACAAA ATTACAACGT CGTGAAAAGG 151 CATCTCGAAC CACGCCACAA CCAGTGCACA CGTGATTTGT TGACGCATTT 201 TGGACTTCGT CAAGATCGTG ATTTGGGTTA CATAAATGTA CACCCCGTAT 251 TTAAGAAAAT AACCTTATTA AATATTGCGC CAAAATACTA CGCGTTATGA 301 TACTATTAGG GTAGGCTTGT GAATTTTACT AAATCGCCCA TCTCGGAATC 351 TAGGTATTIT CTTATATTAA AAAAAATAAG ATGGGGGCCT GCAATTITTT 401 ATTATTTAAT ATTTATTTAT TTTTTAGCGA AGATCCCTCC CTTATTTTAT 451 GAATACCCTT TAATGACTAC ATCTTTATTA TTACTAAGTT TGTCTATAAT 501 TATGAAGTCA ATCTCTACAT ACATAAAAAT AACATATTAA TTACTAATTT 551 AAAACAAATA TTAATGGAAA GTAATATTAC TAAAATTATA ATTACAAACA 601 ACATGGAATT GTCACAAAAT AAAAAATAAA AACTAATTAT CCCATAGTTG 651 GATTAAAATT CATATTGTTA GTATGACTTA AGCTT

FIG. 3. (continued)

p\$217-1 (SEQ ID NO:11)

1 AAGCTTTAAA AGGAAGAGAG CCACAATTTT CTTTGACCTT CCTTCTCCC 51 TAGCCACTAA GATATACAGT ACTGGTCAAA AAGAGCATAT TTATAGCTCA 101 AAATTTTGCC TTTTTCTGTT GTAAACGTGA TTGTTTCTTA CTTGGATTCT 151 TGTTCTATAT ATTTACGGGA GAAAAGAGCA ATTTGCATGC TCCTAAATCT 201 TITATITTCT GGTGAAAAAT TGGTCTTTAA TTGGCTGGGA ATTATTTTT 251 AGATGCTACA ACCTTGACAA ACACCTAAGA ATATTTTAGT GACAATGGCT 301 TGTTCTTTGA GTACTGGTTT TTCTGTTTCT GGTCCCTGTT TCAACGCCAC 351 AGCCAAAGAG TCTCGTCGTC ATTGCCCTTC GATTGGCACT CTGCAACTTA 401 AAGATTTAGC ATCCAGAGAA TTTCTAGGCA AACCCTTGGA TTATGCATCA 451 GATCATATTG GTACTAACCA TTGGAATGTT GAACGACTTT CTGTATGTAA 501 ATCTCTGATA CATTTGCTTC TGTGTTTTATA CTTGGTGTTT TCATGTTTTC 551 ATTCTTGTTT TAAATTTTTC GAGATCAAAT CATTTATAAG TATTTATTCT 601 AATGATTTTA GGCACAAGTA TCAATCGCTG CTCAGAGATG GTGGGAGAAG 651 ACCCTTAAAC CCAACATGGT AGAGATCAAT TCAGCAACAC AACTTGTTGA 701 TTCATTATTA AAAGCTGGTA ATAGATTGGT CATAATTGAC TTCTTCTCTC 751 CTAGCTGTAG AGGTTTCAAG ACTTTACATC CTAAGGTAAG ATATATAGCA 801 ATCCCCTAAA AAAAAAAAA AAAAAAAAA AAACCAACAA CTACATCGTA 851 ATCCTAAGCA AGTTAGGGTT AACTATATGA ATCATCACTA GACGGATCC

pS218 (SEQ ID NO:12)

1 AAGCTTAACT TTACTCACAT TGCTTTCTTT AGGGAAGCGT CTTCTTAAAT 51 GACCATCCTC TAAATTTCTC ATGAATCTTC TTCTGTTGTC CACTCTGTTA 101 TCGCTGAAAC GAAATCTGAA ATTGTCATGA TGCTGACTAT TATCCAATCA 151 CTCAGTCTCT AATTCATATT TAGATTATCT TGTTCACCAG CCCATACTGA 201 TTTTTATTGT TTTGGGGTCT AACTTTTCCT TCCGGTAGTC GGTTGGAGTC 251 ATGAACTTAT TTCTTGAAAT GAGGATATGA CTTTATGGCC TATACTCTTT 301 TGGTGTCTCA AGGCCTGTCA CCTCTCATCT TTTCCTTCAA TTGACTATAG 351 ACTCTGTAAT ACTGTCATCT TTGGGATCTA CCGTTGTCCT CCATGTATCA 401 TATCTTACTC ATAATGCTTC ATTAACTATT TTCTTATTTC CCGCTAACAT 451 TTATGTCTAT CACTITATTC TGAAAACTCG AACAAGACAT TCTTTTCGTT 501 TTAGATCCCC TTTGCTCCAT CCAGTGGTTC TTCGGGGGAC TTAACGTTCT 551 CGCTCTCCTA GGGAGGCGAG CCACACTAAG GTAATATTTA TCCCTTCTAG 601 GCTTTCCGTG CCTATCTTCT GAGATATTTT TTTCATGCTA ATATTCACAT 651 CTAATTGTAA TTTTCTAGAG TGCGCCATCT GGGTGCCTCA CAAGAAGAGC 751 TCCATTCACC ATTTTGGGTT ACTCTAACCT CAGTCGGATA CTAATATCCT 801 GTCATTTTAT TAAACTACAC ATGTTAGCCC CCAATAGGAT ATAACTAAGA 851 TGGGTGTGGC CAATTCTACA TACATCTGTT ACTGTTGAAA GTAAGTCGCA 901 ATGCTTTTAT TTTTCTGCCG GAGTTGAAAA TACCGATAAT CTATATTAAC 951 TGGGTACCTC GTACCCTTCT CATCTTTCTC CTTTTACTTG TTGAAGCTT

FIG. 3. (continued)

pS220-1 (SEO ID NO:13)

1 AAGCTTGAAA AAGAAGAATT AAGGCTTGCT TTCTTAATTT TTAAAAAATA 51 AAAATTATTT TGAACTATCT ATACTATATT AAAAGCACGA AAACCCTATC 101 GAAATGTCGT TCGCCTTTTT TACCCTTTAA AAATAATTTT ACATTAGACA 151 AAATAGTCAT TITACTATIT TICCTAATAT ATAGGATTIT AAAATTAATT 201 TAACTITGGC TATTAAACAT TITCTTATAA CTTGAAATAT GTAAAACTCC 251 TAATATTTAG AAATTTAATT AACATAACCA AGGATTTTTA TATCGGTAAT 301 AACTCTAATA TGGTATCCAA ATCAGTCTAG AACTCTCTTA CCTCTAATAA 351 GTAAAAGTAC TTCTAATAAA TTCATATACT TTTTCTCTCT TCTCCGATCT 401 CTCTTTGCTC TTCTTTTTAT GTATCCTTTC CTTTCTAATA GCCTTTTATG 451 AGAAGTAAAC TTTTAGGGTT GGCCCCCCCT CCCCCCACAA TTATATAGTT 501 TCTTACTCAG TTGTTGGAAT ATAATTCAAA TTCTTAAATA ATTGACGGTG 551 ACATTGAGTT TTACTTTGTG GAAGAGAATT AGATTCTCGT GTTAGTAAAA 601 TCGGTTAGTA ATTGATGATG CATTATTTT ACTCTATAAT AGAGATGCAA 651 TITTATTITT GCATTITGGG ATCAAATTGT AATGCAGTCA TATATTGATT 701 TCATAAATGT TTGGGATATT GTTGGTTATT TAACTAGAAA TAGACTTCTT 751 ATTTCATATT TATTGTTAAA ATCCTTTATT GGAGATGAAT TATTTGTTCA 801 CCGATTAGAA GTTGATAGTC GCTTTTGTTT TAGAAGAAAT TTTACCGTAG 851 ACCAAGTTAA GGAGTTTTAG AAGCACTTTG CATGGGAGCA TTAGTGTATG 901 TTATGGCTTT ATCAAATATA GGTTTTGAAG ATTCAGAGAG CCAAGAAAAG 951 CTAGAACCCA AGAACTAGGA AGTTAGAGTA ATTCACAATA CCATAACGTG 1001 ATATAAAACT TTTTATTGTA ACTCAAATCG GTAATATTTT TTGCTTTAGT 1051 CTTAATCGAT AAATTATTTT TTTATATTGA TTAGTTATAG GAGGCTCACA 1101 AAGTTGGGAA TAATTAAAAT ATCATATTTT GTATTTGAAC AATTTATGAA 1151 ATAGTAATTG GTAAAAAATC ACTTTAAATT TTTATCCTAT ATCCAGAAGG 1201 ATTATGGTGT CTGGCATAGT TGTTTGGAAG ATTTGAATCA GGGTAAAAGT 1251 ATGTTGTAAT TTITATTTTG TTATAGGCAT TTTTTGTGCT TGATTGTTTT 1301 GTTGTCATTA TATTTTATTA TITGGAAGTG TATATATATG TITGATTAAA 1351 ATATAGATAA TCAATTTTAT AAGAAATTTG CAACAATTAC ACAAGGATAA 1401 AGTCTACAAT ATGCGAGTAA AATTTGATTG AACCTAGGAT GTCATATTTA 1451 ATGCATATTT TATTTCAATG TGTTTATTAT ACATCTATTG TATTATATG

FIG. 3. (continued)

CLONE %AT BINDING GRAPHIC REPRESENTATION STRENGTH UU U UUUUUROU ATA TAT AT T pS116-1 76.6 90 (SEQ ID NO: 5) UORR AT RB7-6 73.2 80 (SEQ ID NO: 20) TATTITA A pS211-1 71.7 (SEQ ID NO: 10) URO RU TAT AT TT pS220-1 73.1 70 (SEQ ID NO: 13) UUO U AA A pS226-1 77.1 60 (SEQ ID NO: 9) pS205-2 71.4 (SEQ ID NO: 8) OU TT Ps217-1 65.2 40 (SEQ ID NO: 11) FIG. 4.

CLONE %A+T BINDING STRENGTH

GRAPHIC REPRESENTATION



